JOURNAL OF MULTIVARIATE ANALYSIS 36, 18–34 (1991)

Improved Estimation for a Model Arising in Reliability and Competing Risks

EDSEL A. PEÑA

Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio 43403-0221

Communicated by the Editors

Let $(Z_1, \mathbf{M}_1), ..., (Z_n, \mathbf{M}_n)$ be independent and identically distributed $1 \times (p+1)$ random vectors from the exponential-multinomial distribution which has density function $f(z, \mathbf{m} \mid \theta) = \lambda \exp(-\lambda z) \prod_{j=1}^{p} (\theta_j/\lambda)^{m_j}$ for $z > 0$ and $\mathbf{m} = (m_1, ..., m_p)$ with $m_j \in \{0, 1\}$ and $\mathbf{m}\mathbf{1}_p = 1$, and where $\mathbf{1}_k$ denotes a $k \times 1$ vector of 1's. The parameter $\theta = (\theta_1, ..., \theta_p)$ has $\theta_j > 0$ and $\lambda = \theta \mathbf{1}_p$. This density function arises by observing a series system or a competing risks model with p sources of failure with the lifetime of the ith component or source of failure being exponential with mean $1/\theta_i$, and where the random variable Z denotes system lifetime, while the ith component of $\mathbf{M}$ is a binary random variable denoting whether the ith component failed. It can also arise from the Marshall–Olkin multivariate exponential distribution. The problem of estimating θ with respect to the quadratic loss function $L(\mathbf{a}, \theta) = \|\mathbf{a} - \theta\|^2/\|\theta\|^2$, where $\|\mathbf{v}\|^2 = \mathbf{v}\mathbf{v}'$ for any $1 \times k$ vector $\mathbf{v}$, is considered. Equivariant estimators are characterized, and it is shown that any estimator of form cN/T, where $T = \sum_{i=1}^{n} Z_i$ and $\mathbf{N} = \sum_{i=1}^{n} \mathbf{M}_i$, is inadmissible whenever $c < (n-2)/(n+p-1)$ or $c > (n-2)/n$. Since the maximum likelihood and uniformly minimum variance unbiased estimators correspond to $c\mathbf{N}/T$ with $c = 1$ and $c = (n-1)/n$, respectively, then they are inadmissible. An adaptive estimator, which possesses a self-consistent property, is developed and a second-order approximation to its risk function derived. It is shown that this adaptive estimator is preferable to the estimators $c\mathbf{N}/T$ with $c = (n-2)/(n+p-1)$ and $c = (n-2)/n$. The applicability of the results to the Marshall–Olkin distribution is also indicated.

1. Problem and Setting

Consider a series system with p $(p > 1)$ components or a competing risks model with p sources of failures, and let Y_i $(i = 1, ..., p)$ denote the failure

Received November 10, 1989; revised June 5, 1990.

AMS 1980 subject classifications: primary 62H12; secondary 62C15, 62N05.

Key words and phrases: equivariant estimator, exponential-multinomial distribution, inadmissibility, invariance, Marshall–Olkin multivariate exponential, self-consistent estimator.

0047-259X/91 $3.00

time of the ith component or source of failure. Upon system failure the observed random vector is $(Z, M_1, ..., M_p)$, where $Z = \min(Y_1, ..., Y_p)$ and $M_i = I(Z = Y_i)$ $(i = 1, ..., p)$ with $I(A)$ denoting the indicator function of event A. If the Y_i's are independent and Y_i has exponential distribution with mean $1/\theta_i$, it is easy to see that the joint density function of $(Z, M_1, ..., M_p)$ is

$$f(z, \mathbf{m} \mid \theta) = \lambda \exp(-\lambda z) \prod_{j=1}^{p} \left(\frac{\theta_j}{\lambda}\right)^{m_j} \tag{1.1}$$

for $z > 0$ and $\mathbf{m} = (m_1, ..., m_p)$, where $m_j \in \{0, 1\}$ with $\mathbf{m}\mathbf{1}_p = 1$, $\lambda = \theta \mathbf{1}_p$ with $\theta = (\theta_1, ..., \theta_p)$, and $\mathbf{1}_k$ denotes a $k \times 1$ vector of 1's for any positive integer k. As will be seen in Section 5, (1.1) also arises from the Marshall–Olkin [6] multivariate exponential (MVE) distribution under a series sampling scheme. We shall refer to (1.1) as the exponential-multinomial density with parameters p and θ, and write this as $\mathrm{EM}(p, \theta)$. Typically the parameter p is known and inferential problems pertain to θ.

Let $(Z_i, \mathbf{M}_i)$ $(i = 1, ..., n)$ be independent and identically distributed (i.i.d.) $1 \times (p+1)$ random vectors from $\mathrm{EM}(p, \theta)$, where θ belongs to the parameter space $\Theta = \{\theta \mid \theta = (\theta_1, ..., \theta_p),\ \theta_j > 0\ (j = 1, ..., p)\}$, but is otherwise unknown. Letting $\|\mathbf{v}\|^2 = \mathbf{v}\mathbf{v}' = \sum_{j=1}^{k} v_j^2$ for any $1 \times k$ vector $\mathbf{v}$, with prime denoting transpose, we consider the estimation of θ with respect to the quadratic loss function

$$L(\mathbf{a}, \theta) = \frac{\|\mathbf{a} - \theta\|^2}{\|\theta\|^2}, \tag{1.2}$$

where $\mathbf{a}$ belongs to the action space $\mathscr{A} = \{\mathbf{a} \mid \mathbf{a} = (a_1, ..., a_p),\ a_j \geqslant 0$ $(j = 1, ..., p)\}$. We note that risk functions arising from (1.2) yield risk functions from the more popular loss $\|\mathbf{a} - \theta\|^2$, hence inadmissibility results based on (1.2) are equivalent to inadmissibility results based on the latter loss function. We adopt (1.2) for notational convenience and in anticipation of future minimaxity studies. In the context of the MVE distribution several papers dealt with statistical inferences about θ, although it seems that none of them considered a decision-theoretic approach. Arnold [1] obtained unbiased estimators of θ, while Bemis, Bain, and Higgins [2] presented maximum likelihood (ML) estimators. Proschan and Sullo [9, 10] also derived ML estimators of θ under the series, parallel, and cause-identifiable sampling schemes. Shamseldin and Press [11] approached the estimation problem via a Bayesian framework with their prior on θ assuming independence of components, while Peña and Gupta [8] considered priors which can model prior dependence among the components.

Now, by factorization theorem the $1 \times (p+1)$ random vector $(T, \mathbf{N})$ is sufficient for θ, where $(T, \mathbf{N}) = \mathbf{1}_n'(\mathbf{Z}, \mathbf{M})$ and $(\mathbf{Z}, \mathbf{M})$ is the $n \times (p+1)$ matrix with ith row $(Z_i, \mathbf{M}_i)$. By the Sufficiency Principle the inferential problem can be reduced by considering $(T, \mathbf{N})$ as the observable random vector. The relevant sample space is therefore $\mathscr{X} = \{(t, \mathbf{n}) \mid t > 0,\ \mathbf{n} = (n_1, ..., n_p),\ n_j \in \mathbb{Z},\ \mathbf{n}1_p = n\}$. We denote the space of estimators of θ by $\mathscr{D} = \{\delta \mid \delta: \mathscr{X} \to \mathscr{A}$, where δ is a measurable map$\}$. For $\delta \in \mathscr{D}$ the risk function is $R(\delta, \theta) = E_\theta[L(\delta, \theta)]$, and $\delta \in \mathscr{D}$ is inadmissible if there exists a $\delta^* \in \mathscr{D}$ such that $R(\delta^*, \theta) \leqslant R(\delta, \theta)$ for every $\theta \in \Theta$ and with strict inequality for some $\theta \in \Theta$. From (1.1) and the definition of T and $\mathbf{N}$ it is easy to see that

(i) T and $\mathbf{N}$ are independent;

(ii) T has a gamma distribution with parameters n and λ;

(iii) $\mathbf{N}$ has a singular multinomial distribution with parameters n and $\beta = (\beta_1, ..., \beta_p)$, (1.3)

where $\beta = \theta/\lambda = (\theta_1, ..., \theta_p)/\lambda$. Without going into details (cf. Peña [7]), it can then be established that the estimation problem specified by the triplet $(\Theta, \mathscr{A}, L)$ and the distributions P_θ over $\mathscr{X}$ specified in (1.3) is invariant (cf. Ferguson [4]) under the group of transformations $\mathscr{G} = \{g \mid g: \mathscr{X} \to \mathscr{X}$ with $g(t, \mathbf{n}) = (ct, \mathbf{n})$ for some $c > 0\}$. Letting $\mathscr{E} \subset \mathscr{D}$ denote the subclass of equivariant estimators (cf. Lehmann [5]), we find that $\mathscr{E} = \{\delta \mid \delta \in \mathscr{D},\ \delta(T, \mathbf{N}) = h(\mathbf{N})/T$ for some measurable map $h: \mathbb{R}_p \to \mathbb{R}_p\}$. Invoking the Invariance Principle it suffices to limit attention to the subclass $\mathscr{E}$. The estimators considered in this paper are therefore in $\mathscr{E}$. As will be seen in Section 2 the ML and uniformly minimum variance (UMVU) estimators of θ belong in $\mathscr{E}$, but it will also be shown that these popular estimators are inadmissible. In Section 3 we attempt to improve on the estimators that dominate the ML and UMVU estimators by considering an adaptive estimator of the form $g(\mathbf{N})\mathbf{N}/T$, where $g: \mathbb{R}_p \to \mathbb{R}$. Exact expression for the risk of this new estimator is difficult to obtain, so a second-order approximation is presented. Algebraic comparisons of the risk functions are then performed in Section 4. Finally in Section 5 we indicate how the new estimator can be applied in estimating the parameter of the MVE distribution.

2. Inadmissibility of ML and UMVU Estimators

We start by considering the subclass $\mathscr{E}_1 \subset \mathscr{E}$ defined by $\mathscr{E}_1 = \{\delta_c \mid \delta_c \in \mathscr{E},\ \delta_c(T, \mathbf{N}) = c\mathbf{N}/T$, where $c \geqslant 0\}$.

LEMMA 2.1. *Let $n > 2$ and $\delta_c \in \mathscr{E}_1$. Then*

$$R(\delta_c, \theta) = c^2 \left[\frac{n(n+\rho^2-1)}{(n-1)(n-2)}\right] - 2c\left(\frac{n}{n-1}\right) + 1, \tag{2.1}$$

where $\rho^2 = 1/\|\beta\|^2 = \lambda^2/\|\theta\|^2$.

Proof. From (1.3) and well-known results about the gamma distribution and multinomial distribution, we obtain

$$E_\theta\left(\frac{1}{T}\right) = \frac{\lambda}{n-1} \quad \text{and} \quad E_\theta\left(\frac{1}{T^2}\right) = \frac{\lambda^2}{(n-1)(n-2)} \tag{2.2}$$

and

$$E_\theta(\mathbf{N}) = n\beta \quad \text{and} \quad \mathrm{Cov}_\theta(\mathbf{N}, \mathbf{N}) = n\Sigma, \tag{2.3}$$

where Σ is the $p \times p$ matrix with diagonal elements $\beta_i(1-\beta_i)$ $(i = 1, ..., p)$ and off-diagonal elements $-\beta_i\beta_j$ $(i \neq j,\ i, j = 1, ..., p)$. From (2.3) we also have

$$\begin{aligned} E_\theta(\|\mathbf{N}\|^2) &= E_\theta\left(\sum_{j=1}^{p} N_j^2\right) = \sum_{j=1}^{p} \{\mathrm{Var}_\theta(N_j) + [E_\theta(N_j)]^2\} \\ &= \sum_{j=1}^{p} [n\beta_j(1-\beta_j) + (n\beta_j)^2] = n + n(n-1)\,\|\beta\|^2. \end{aligned} \tag{2.4}$$

The risk function of δ_c is now given by

$$\begin{aligned} R(\delta_c, \theta) &= E_\theta\left(\frac{\|c(\mathbf{N}/T) - \theta\|^2}{\|\theta\|^2}\right) = \frac{c^2 E_\theta(\|\mathbf{N}/T\|^2) - 2cE_\theta(\mathbf{N}/T)\theta' + \|\theta\|^2}{\|\theta\|^2} \\ &= \frac{c^2 E_\theta(\|\mathbf{N}\|^2)\, E_\theta(1/T^2) - 2cE_\theta(\mathbf{N})\, E_\theta(1/T)\theta' + \|\theta\|^2}{\|\theta\|^2}, \end{aligned} \tag{2.5}$$

where we used the independence of $\mathbf{N}$ and T (see 1.3(i)). Substituting the expressions in (2.2), (2.3), and (2.4) in (2.5), and then simplifying, we obtain (2.1). ∎

COROLLARY 2.1. *Let $n > 2$ and set $c_0 = (n-2)/(n+\rho^2-1)$. Then, for any nonnegative real numbers c_1 and c_2, $R(\delta_{c_1}, \theta) < R(\delta_{c_2}, \theta)$ for every $\theta \in \Theta$ whenever $c_0 \leqslant c_1 < c_2$ or $c_0 \geqslant c_1 > c_2$.*

Proof. From Lemma 2.1, since $R(\delta_c, \theta)$ is parabolic in c, then $R(\delta_c, \theta)$ attains its minimum at the vertex $c = c_0$. ∎

THEOREM 2.1. *Let $n > 2$. Then any estimator $\delta_c \in \mathscr{E}_1$ with $c > (n-2)/n$ or $c < (n-2)/(n+p-1)$ is inadmissible.*

Proof. From the definition of ρ^2 in Lemma 2.1 and by Cauchy–Schwarz inequality, $\lambda^2 = (\theta \mathbf{1}_p)^2 \leqslant \|\theta\|^2 p$, implying that $\rho^2 \leqslant p$. Since $\rho^2 > 1$, then

$$1 < \rho^2 \leqslant p. \tag{2.6}$$

By the definition of c_0 in Corollary 2.1 and using (2.6), we obtain the inequalities

$$\frac{n-2}{n+p-1} \leqslant c_0 \equiv \frac{n-2}{n+\rho^2-1} < \frac{n-2}{n}. \tag{2.7}$$

Applying Corollary 2.1 and by (2.7), $\delta_c \in \mathscr{E}_1$ with $c > (n-2)/n$ is then dominated by the estimator

$$\delta_1^* \equiv \delta_{(n-2)/n} = \left(\frac{n-2}{n}\right)\frac{\mathbf{N}}{T}, \tag{2.8}$$

while $\delta_c \in \mathscr{E}_1$ with $c < (n-2)/(n+p-1)$ is then dominated by the estimator

$$\delta_2^* \equiv \delta_{(n-2)/(n+p-1)} = \left(\frac{n-2}{n+p-1}\right)\frac{\mathbf{N}}{T}. \quad \blacksquare \tag{2.9}$$

COROLLARY 2.2. *The ML and UMVU estimators are inadmissible.*

Proof. The ML estimator of θ is $\delta_1 \in \mathscr{E}_1$, while the UMVU estimator of θ is $\delta_{(n-1)/n} \in \mathscr{E}_1$. Since $(n-2)/n < (n-1)/n < 1$ then these estimators are inadmissible by Theorem 2.1. $\blacksquare$

3. A BETTER ADAPTIVE ESTIMATOR

At this juncture a question arises whether the estimators (2.8) and (2.9) are themselves admissible. Clearly if ρ^2 is known with $\rho^2 < p$, then by Corollary 2.1, $\delta_{c_0} \in \mathscr{E}_1$ dominates (2.8) and (2.9). However, if ρ^2 is unknown, which is usually the case, δ_{c_0} is no longer an estimator. A related question is which of (2.8) and (2.9) is a better estimator. From the definition of ρ^2 we note that it will be near p (see (2.6)) when the components of θ are approximately equal, while it will be near one when a single component of θ dominates the other components. Thus by construction of (2.8) and (2.9), we expect (2.8) to be better than (2.9) under the latter situation, while (2.9) will be better than (2.8) in the other situation. Neither of them, however, will dominate the other for all values of θ, and the choice of which estimator to use will depend on the value of the unknown parameter ρ^2.

As pointed out above, if ρ^2 is unknown then δ_{c_0} is no longer an estimator, since c_0 depends on ρ^2. An idea then is to estimate ρ^2 by say $\hat{\rho}^2$ and then estimate c_0 by $\hat{c}_0$, where the ρ^2 in c_0 (see Corollary 2.1) is replaced by $\hat{\rho}^2$. If we could find a good estimator of ρ^2, then there is hope that the adaptive estimator $\delta_{\hat{c}_0}$ will be preferable to either (2.8) or (2.9). Clearly such an estimator will not be in $\mathscr{E}_1$ and consequently its risk function cannot be obtained via Lemma 2.1.

There are several possible estimators of ρ^2, but the most appealing is the estimator

$$\hat{\rho}^2 = \frac{n^2}{\|\mathbf{N}\|^2}. \tag{3.1}$$

It is the ML estimator of ρ^2 by invariance of ML estimators, and it satisfies a "self-consistency" property akin to that of the product-limit estimator [3]. For if at the jth stage of the iteration δ_j is the estimator of θ, then we could estimate ρ^2 by $(\delta_j \mathbf{1}_p)^2/\|\delta_j\|^2$, and consequently obtain an updated estimator of θ given by

$$\delta_{j+1} = \left[\frac{(n-2)}{n + (\delta_j \mathbf{1}_p)^2/\|\delta_j\|^2 - 1}\right] \frac{\mathbf{N}}{T}.$$

If this iteration converges then the final estimator should satisfy the identity

$$\delta = \left[\frac{(n-2)}{n + (\delta \mathbf{1}_p)^2/\|\delta\|^2 - 1}\right] \frac{\mathbf{N}}{T}; \tag{3.2}$$

and from this δ we obtain our final estimator of ρ^2 given by $(\delta \mathbf{1}_p)^2/\|\delta\|^2$. But this is precisely (3.1). Our adaptive estimator is therefore

$$\delta_3^* = \left[\frac{(n-2)\,\|\mathbf{N}\|^2}{n^2 + (n-1)\,\|\mathbf{N}\|^2}\right] \frac{\mathbf{N}}{T}. \tag{3.3}$$

A closed-form expression for the risk function of (3.3) is difficult to obtain due to the form of the shrinkage factor which is now random. We are able, however, to obtain a second-order approximation given below.

THEOREM 3.1. *$R(\delta_3^*, \theta) = \rho^2/(n + \rho^2 - 1) + (1 + 4\rho^2\gamma^2)/(n + \rho^2 - 1)^2 + O(n^{-5/2})$ as $n \to \infty$, where $\gamma^2 = \beta\Sigma\beta'/(\|\beta\|^2)^2$.*

Proof. We defer the proof to the Appendix.

The function of θ given by γ^2 in Theorem 3.1 can also be expressed as

$$\gamma^2=\rho^4\left(\sum_{i=1}^{p}\beta_i^3\right)-1=\lambda\left(\sum_{i=1}^{p}\frac{\theta_i^3}{(\|\theta\|^2)^2}\right)-1.$$

This follows easily from $\beta\Sigma\beta'=\sum_{i=1}^{p}\beta_i^3-(\|\beta\|^2)^2$. Note also that if $\mathbf{V}$ is a $1\times p$ multinomial random vector with parameters 1 and β, then $\gamma^2=\mathrm{Var}(\beta\mathbf{V}')/[E(\beta\mathbf{V}')]^2=[CV(\beta\mathbf{V}')]^2$. Thus, γ^2 will be near zero whenever either the components of β or θ are near each other or when one component dominates the others. This behavior contrasts with that of ρ^2 which becomes small (near one) only when one of the components dominates the others.

4. Comparison of Risks

Straightforward substitutions of the appropriate c in $R(\delta_c, \theta)$ of Lemma 2.1 and then simplifying yield the risk functions of (2.8) and (2.9) to be, respectively,

$$R(\delta_1^*, \theta)=\frac{(n-2)\rho^2+2}{n(n-1)} \tag{4.1}$$

and

$$R(\delta_2^*, \theta)=\frac{n(n-2)\rho^2+(n-1)(1+p^2)+2p}{(n-1)(n+p-1)^2}. \tag{4.2}$$

To illustrate more clearly the behavior of these functions as ρ^2 varies over its range $(1, p]$ (see (2.6)), (4.1) and (4.2) could be re-expressed as

$$R(\delta_1^*, \theta)=\frac{1}{(n-1)}\left[1+\left(\frac{n-2}{n}\right)(\rho^2-1)\right] \tag{4.1'}$$

and

$$R(\delta_2^*, \theta)=\left[\frac{1+(n-1)p}{(n-1)(n+p-1)}\right]\left\{1-\frac{n(n-2)(p-\rho^2)}{[1+(n-1)p](n+p-1)}\right\}. \tag{4.2'}$$

Thus we see from (4.1′) and (4.2′) that as ρ^2 varies from 1 to p, (4.1) and (4.2) both increase. It is easy to show then that there exist $\varepsilon_1>0$ and $\varepsilon_2>0$, both depending on n and p, such that for $\rho^2\in(1, 1+\varepsilon_1)$ we have $(4.1)<(4.2)$, and for $\rho^2\in(p-\varepsilon_2, p]$ we have $(4.2)<(4.1)$. This confirms formally the observation made in the first paragraph of the preceding section concerning the preferability of (2.8) and (2.9).

On the other hand, using the observation in the last paragraph of the preceding section that $\gamma^2 \to 0$ as $\rho^2 \to 1$ or $\rho^2 \to p$, we obtain from Theorem 3.1 that, as $n \to \infty$,

$$R(\delta_3^*, \theta) \to \frac{1}{(n-1)}\left(1 - \frac{1}{n^2}\right) + O(n^{-5/2}) \qquad \text{as} \quad \rho^2 \to 1 \tag{4.3}$$

and

$$R(\delta_3^*, \theta) \to \frac{1}{(n+p-1)}\left[p + \frac{1}{(n+p-1)}\right] + O(n^{-5/2}) \qquad \text{as} \quad \rho^2 \to p. \tag{4.4}$$

From (4.1′) and (4.3), and (4.2′) and (4.4), it then follows that

THEOREM 4.1. $\lim_{\rho^2 \to 1}[R(\delta_1^*, \theta) - R(\delta_3^*, \theta)] = O(n^{-5/2})$ *and* $\lim_{\rho^2 \to p}$ $[R(\delta_2^*, \theta) - R(\delta_3^*, \theta)] = O(n^{-5/2})$, *as* $n \to \infty$.

These results show that (3.3) is risk-equivalent to (2.8) and (2.9) up to terms of order n^{-2} in the cases where (2.8) and (2.9) are expected to perform well. More generally, we have the following results.

THEOREM 4.2. *For large n, and ignoring* $O(n^{-5/2})$ *terms,*

(i) $R(\delta_3^*, \theta) < R(\delta_1^*, \theta)$ *iff* $(\rho^2 - 1) > 2\gamma(\sqrt{\gamma^2 + 1} + \gamma)$, *and*

(ii) $R(\delta_3^*, \theta) < R(\delta_2^*, \theta)$ *iff* $(p - \rho^2) > 2\gamma(\sqrt{\gamma^2 + p} - \gamma)$.

Proof. We will just present the proof of (i), since the proof of (ii) is analogous, although longer. From (4.1) and Theorem 3.1 we have

$$\begin{aligned} R(\delta_1^*, \theta) - R(\delta_3^*, \theta) &= \rho^2\left\{\frac{n-2}{n(n-1)} - \frac{1}{(n+\rho^2-1)}\right\} \\ &\quad + \left\{\frac{2}{n(n-1)} - \frac{1 + 4\rho^2\gamma^2}{(n+\rho^2-1)^2}\right\} + O(n^{-5/2}), \end{aligned}$$

which after absorbing terms of order $O(n^{-3})$ into $O(n^{-5/2})$ simplifies to

$$\frac{\rho^2(\rho^2-2)}{(n-1)(n+\rho^2-1)} + \frac{1 - 4\rho^2\gamma^2}{(n+\rho^2-1)^2} + O(n^{-5/2}).$$

Ignoring the $O(n^{-5/2})$ term, the first two terms will have sum greater than zero iff

$$\rho^4(\rho^2 - 2) + (n-1)[\rho^2(\rho^2 - 2) + 1 - 4\rho^2\gamma^2] > 0. \tag{4.5}$$

Since $\rho^2(\rho^2-2)+1-4\rho^2\gamma^2>0$ iff $\rho^2>1+2[\gamma^2+\sqrt{\gamma^2(1+\gamma^2)}]$, then (4.5) holds iff

$$n\{<,>\}1-\frac{\rho^4(\rho^2-2)}{\rho^2(\rho^2-2)+1-4\rho^2\gamma^2}$$

with the appropriate inequality chosen according to whether

$$\rho^2\{<,>\}1+2[\gamma^2+\sqrt{\gamma^2(1+\gamma^2)}],$$

respectively. For large n the first possibility will not be satisfied and the $O(n^{-5/2})$ term is negligible. Thus, (i) is proved. ∎

It is interesting to note the similarities of the conditions for which (3.3) dominates (2.8) and (2.9) up to terms of order n^{-2}. From (2.6), (ρ^2-1)

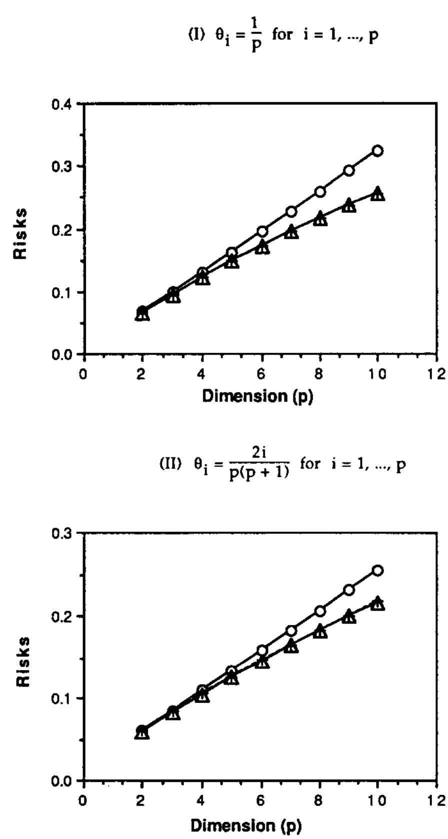
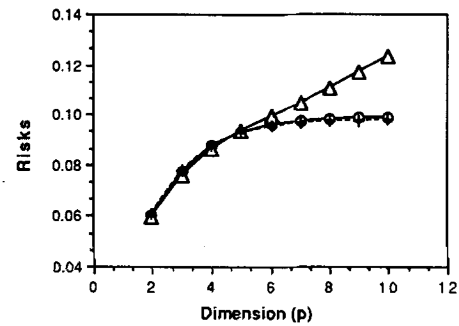


FIG. 1. Risks of estimators δ_1^*, δ_2^*, and δ_3^* for $n=30$ and $p=2, 3, ..., 10$ under four types of θ-vectors. The risks for δ_3^* are up to terms of order n^{-2}. $\delta_1^*=\bigcirc$, solid lines; $\delta_2^*=\triangle$, solid lines; and $\delta_3^*=+$, dashed lines.

(III) $\theta_i = \dfrac{2^{i-1}}{2^p - 1}$ for $i = 1, \ldots, p$



(IV) $\theta_i = \dfrac{1}{11(p-1)}$ for $i = 1, \ldots, p-1$; $\theta_p = \dfrac{10}{11}$

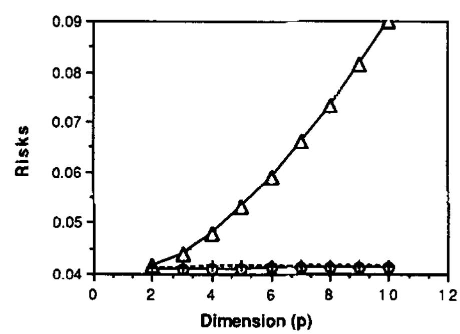


FIG. 1—*Continued*

and $(p-\rho^2)$ are the distances of ρ^2 from its lower and upper bounds, respectively; and we recall that (2.8) performs well if ρ^2 achieves its lower bound, while (2.9) performs well when ρ^2 achieves its upper bound. Thus Theorem 4.2 confirms formally that (3.3) dominates (2.8) when ρ^2 is far from 1, and (3.3) dominates (2.9) when ρ^2 is far from p. Nevertheless, Theorem 4.2 also shows that (3.3) does not dominate (2.8) or (2.9) for all values of θ. When ρ^2 is near 1, (2.8) will be better than (3.3), and this can be intuitively explained by the fact that the constant estimator 1 is better than the estimator $n^2/\|\mathbf{N}\|^2$ for estimating ρ^2. Similarly, when ρ^2 is near p, the constant estimator p is better than $n^2/\|\mathbf{N}\|^2$ for estimating ρ^2, so (2.9) is better than (3.3). We deduce from Theorem 4.1, however, that the differences in risks between (2.8) and (3.3) and (2.9) and (3.3), under these extreme situations will be negligible. To see this graphically, we plotted the risks of (2.8), (2.9), and (3.3) for $n=30$ and $p=2, 3, \ldots, 10$ for four types of

θ-vectors given by: (I) $\theta_i = 1/p$ for $i = 1, ..., p$; (II) $\theta_i = 2i/p(p+1)$ for $i = 1, ..., p$; (III) $\theta_i = 2^{i-1}/(2^p - 1)$ for $i = 1, ..., p$; and (IV) $\theta_i = 1/11(p-1)$ for $i = 1, ..., p-1$ and $\theta_p = 10/11$. These types were chosen to achieve different patterns for ρ^2 and γ^2. For instance, for type I, $\rho^2 = p$ and $\gamma^2 = 0$; for type II, ρ^2 still increases with p while $\gamma^2 \neq 0$ but is almost constant with respect to p; for type III, ρ^2 does not increase much when p increases; while type IV has one component dominating the others so ρ^2 and γ^2 are both small. Figure 1 contains the plotted risks. As can be seen from these graphs, (2.8) performs well for type III and type IV θ-vectors, (2.9) performs well for type I and type II θ-vectors, while (3.3) performs well on all four θ-vectors. Though (3.3) is not uniformly better, the differences in risks between (3.3) and whichever is better between (2.8) and (2.9) are negligible.

Going back to Theorem 4.2 one may pose the question: how "large" are the regions in Θ, or equivalently in the (ρ^2, γ^2)-space for which (3.3) is better than (2.8) and (2.9)? To partially answer this question we generated 10,000 θ-vectors for each $p = 2, 3, ..., 10$ according to two models: (i) $\theta_1, ..., \theta_p$ are i.i.d. unit uniform, and (ii) $\theta_1, ..., \theta_p$ are i.i.d. unit exponential. Note by invariance of ρ^2 and γ^2 with respect to a scale change that it

TABLE I

Summary of Simulation Study for Determining the Probabilities that δ_3^* Will Dominate δ_1^* and δ_2^*, Together with the Characteristics of ρ^2 and γ^2 under Two Models for Generating θ-Vectors

	Model for generating the θ-vector									
	$\theta_1, ..., \theta_p$ i.i.d. $U(0,1)$[a]					$\theta_1, ..., \theta_p$ i.i.d. Exp(1)[b]				
p	$\delta_3^* > \delta_1^*$[c]	$\delta_3^* > \delta_2^*$[d]	Mean of ρ^2	Mean of γ^2	Corr of (ρ^2, γ^2)	$\delta_3^* > \delta_1^*$[c]	$\delta_3^* > \delta_2^*$[d]	Mean of ρ^2	Mean of γ^2	Corr of (ρ^2, γ^2)
2	0.5098[e]	0.1489[e]	1.6864	0.0615	−0.5201	0.3516[e]	0.2560[e]	1.5691	0.0704	−0.2806
3	0.7968	0.2258	2.4064	0.0911	−0.6147	0.5569	0.4161	2.1149	0.1170	−0.4246
4	0.9278	0.2959	3.1430	0.1037	−0.6446	0.6996	0.5395	2.6601	0.1519	−0.5173
5	0.9802	0.3602	3.8889	0.1113	−0.6436	0.7910	0.6551	3.1891	0.1802	−0.5603
6	0.9950	0.4292	4.6366	0.1162	−0.6399	0.8599	0.7616	3.7180	0.2015	−0.6057
7	0.9995	0.5139	5.3704	0.1190	−0.6397	0.8977	0.8413	4.2145	0.2240	−0.6265
8	0.9996	0.5894	6.1150	0.1204	−0.6473	0.9293	0.8979	4.7447	0.2392	−0.6550
9	1.0000	0.6574	6.8633	0.1216	−0.6441	0.9510	0.9380	5.2389	0.2526	−0.6572
10	1.0000	0.7204	7.6198	0.1221	−0.6448	0.9705	0.9669	5.7590	0.2654	−0.6687

[a] The IMSL routine RNUN was used to generate uniform variates.
[b] The IMSL routine RNEXP was used to generate exponential variates.
[c] Proportion for which δ_3^* is better than δ_1^*.
[d] Proportion for which δ_3^* is better than δ_2^*.
[e] Maximum standard error is 0.005.

suffices to consider the uniform on (0, 1) and the unit exponential. For each set of 10,000 θ-vectors the proportions for which $(\rho^2-1)>2\gamma[\sqrt{\gamma^2+1}+\gamma]$ and $(p-\rho^2)>2\gamma[\sqrt{\gamma^2+p}-\gamma]$ were determined. Means, variances, and correlations of ρ^2 and γ^2 were also computed. Results of this simulation are summarized in Table I. From this table we see that under both models for generating θ, (3.3) is better than (2.8) with high probability especially when p is large. Estimator (2.9), on the other hand, fares better than (3.3) especially when p is small and $\theta_1, ..., \theta_p$ are i.i.d. uniform. This can be explained by the fact that under this model, the generated θ-vectors tend to have large ρ^2 as can be seen from the means of ρ^2, and this is the situation suited to (2.9). As p becomes larger, however, (3.3) becomes better than (2.9) with high probability under both models. Based on these discussions, it is then safe to conclude that (3.3) is preferable than (2.8) and (2.9), since it performs reasonably well under all circumstances, even though it does not make (2.8) and (2.9) inadmissible.

5. Applications to an MVE Distribution

Following Marshall and Olkin [6] for a fixed integer $k\geqslant 2$, let J denote the set of vectors $\mathbf{e}=(e_1, ..., e_k)$ with $e_i \in \{0, 1\}$ and $\mathbf{e}\neq\mathbf{0}=(0, ..., 0)$. Thus J has 2^k-1 elements, and in order to relate to the notations in the preceding sections, we set $p\equiv 2^k-1$. Furthermore, for notational convenience, we assume that the elements of J are ordered lexicographically and denote these elements by $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_p$. Now, a random vector $\mathbf{X}=(X_1, ..., X_k)$ is said to follow a Marshall–Olkin multivariate exponential (MVE) distribution with parameter $\theta=(\theta_{\mathbf{e}_1}, \theta_{\mathbf{e}_2}, ..., \theta_{\mathbf{e}_p})$, where $\theta_{\mathbf{e}_j}>0$ $(j=1, ..., p)$, if for every $\mathbf{x}=(x_1, ..., x_k)\in\mathbb{R}_k$ with $x_i>0$ $(i=1, ..., k)$,

$$P(X_1>x_1, ..., X_k>x_k)=\exp\left[-\sum_{j=1}^{p}\theta_{\mathbf{e}_j}\max(\mathbf{x}\mathbf{e}_j)\right], \tag{5.1}$$

where, for $\mathbf{e}\in J$, $\max(\mathbf{xe})=\max\{x_i e_i: i=1, ..., k\}$. This distribution has the appealing properties of having exponential marginals and a memoryless property similar to the univariate exponential distribution. Marshall and Olkin [6] have shown that (5.1) could arise in both fatal and non-fatal shock models, and hence it is useful in modelling lifetimes of dependent components in reliability, biometry, and survival analysis. The papers mentioned in Section 1 dealt with inferences about θ.

Suppose now that $\mathbf{X}_1, ..., \mathbf{X}_n$ is a random sample from (5.1) and we want to estimate θ according to the loss function in (1.2). If one is to use directly the vectors $\mathbf{X}_1, ..., \mathbf{X}_n$, difficulties are encountered, since (5.1) is not absolutely continuous with respect to k-dimensional Lebesgue measure. A dominating measure can be constructed (cf. Proschan and Sullo [10]) but the resulting estimators are usually hard to obtain. A simpler solution, proposed by Arnold [1] and which coincides with the series sampling scheme (cf. [9]), is to define for every $\mathbf{e} \in J$,

$$Z_i = \min_{1 \leqslant j \leqslant k} X_{ij} \qquad \text{and} \qquad \delta_{\mathbf{e}, i} = I(Z_i = X_{ij} \text{ for every } j \text{ with } e_j = 1).$$

It is easy to show that the random vector $(Z_i, \delta_{\mathbf{e}, i} : \mathbf{e} \in J)$ follows an $\mathrm{EM}(p, \theta)$ distribution. Letting $(T, N_{\mathbf{e}} : \mathbf{e} \in J)$ with

$$T = \sum_{i=1}^{n} Z_i \qquad \text{and} \qquad N_{\mathbf{e}} = \sum_{i=1}^{n} \delta_{\mathbf{e}, i} \qquad (\mathbf{e} \in J),$$

we are then led into the setting described in Section 1, and θ could be estimated using (3.3). The ML and UMVU estimators of θ under the series sampling scheme were first presented by Arnold [1], although he did not show that these estimators were inadmissible with respect to squared-error loss or the equivalent loss function in (1.2). Note, by the way, that the estimator $\delta_{(n-1)/n} = ((n-1)/n)(\mathbf{N}/T)$ is no longer UMVU when $X_1, ..., X_n$ are observed.

Appendix

For the proof of Theorem 3.1 we need a lemma which gives second-order approximations of moments of functions of $\mathbf{N}$. The proof of this lemma is lengthy and can be found in Peña [7, Lemma 4.1].

LEMMA A.1. *Let $\mathbf{X}$ be a $1 \times p$ random vector with $n\mathbf{X}$ having a singular multinomial distribution with mean vector $n\beta$ and covariance matrix $n\Sigma$. Let $f: \mathbb{R}_p \to \mathbb{R}$ and suppose there exists $g: \mathbb{R} \to \mathbb{R}$ and $h: \mathbb{R} \to \mathbb{R}$ such that*

$$f(\mathbf{x}) = g(\|\mathbf{x}\|^2)\, h(\mathbf{x}\beta') \qquad \forall \mathbf{x}.$$

The functions f, g, and h may depend on n. If $f_i = \partial f / \partial x_i$ and $f_{ij} = \partial^2 f / \partial x_i\, \partial x_j$ $(i, j = 1, ..., p)$ exist and are continuous at $\mathbf{x} = \beta$, and there exists a $B(n)$ such that $|\partial^3 f / \partial x_i\, \partial x_j\, \partial x_k| \leqslant B(n)$ $\forall \mathbf{x}$, then, as $n \to \infty$,

$$\begin{aligned} E[f(\mathbf{X})] = f(\|\beta\|^2) + \frac{1}{n} \Big\{ & g^{(1)}(\|\beta\|^2)\, h(\|\beta\|^2)(1 - \|\beta\|^2) \\ & + \frac{1}{2} [g(\|\beta\|^2)\, h^{(2)}(\|\beta\|^2) + 4g^{(1)}(\|\beta\|^2)\, h^{(1)}(\|\beta\|^2) \\ & + 4g^{(2)}(\|\beta\|^2)\, h(\|\beta\|^2)](\beta\Sigma\beta') \Big\} + B(n)\, O(n^{-3/2}), \end{aligned}$$

where $a^{(j)}$ denotes the jth *derivative of a function* $a\colon \mathbb{R} \to \mathbb{R}$, *and* $b(n) = O(n^k)$ *means that* $b(n)/n^k$ *remains bounded as* $n \to \infty$.

Proof of Theorem 3.1. From (3.3) note that δ_3^* could be expressed as

$$\delta_3^* = \left[\frac{(n-2)\,\|\mathbf{X}\|^2\,\mathbf{X}}{1+(n-1)\,\|\mathbf{X}\|^2}\right]\frac{n}{T},$$

where $\mathbf{X} = \mathbf{N}/n$ has the same distribution as the $\mathbf{X}$ in Lemma A.1. Consequently, by using the independence of $\mathbf{X}$ and T, we have

$$R(\delta_3^*, \theta) = E\left\{\frac{(n-2)^2\,(\|\mathbf{X}\|^2)^3}{[1+(n-2)\,\|\mathbf{X}\|^2]^2}\right\} E\left(\frac{n^2}{T^2\,\|\theta\|^2}\right) - 2E\left\{\frac{(n-2)\,\|\mathbf{X}\|^2\,\mathbf{X}\beta'}{1+(n-2)\,\|\mathbf{X}\|^2}\right\} E\left(\frac{n\lambda}{T\,\|\theta\|^2}\right) + 1. \qquad \text{(A.1)}$$

From (2.2) we obtain

$$E\left(\frac{n^2}{T^2\,\|\theta\|^2}\right) = \frac{n^2\rho^2}{(n-1)(n-2)} \quad \text{and} \quad E\left(\frac{n\lambda}{T\,\|\theta\|^2}\right) = \frac{n\rho^2}{(n-1)}. \qquad \text{(A.2)}$$

Applying Lemma A.1 with $g(y) = (n-2)^2\,y^3/[1+(n-1)\,y]^2$ and $h(y) = 1$ and, using the fact that $1/p \leqslant \|\mathbf{X}\|^2 \leqslant 1$, we obtain

$$\begin{aligned} E&\left\{\frac{(n-2)^2\,(\|\mathbf{X}\|^2)^3}{[1+(n-1)\,\|\mathbf{X}\|^2]^2}\right\} \\ &= \frac{(n-2)^2\,(\|\beta\|^2)^3}{[1+(n-1)\,\|\beta\|^2]^2} \\ &\quad + \frac{1}{n}\left\{\frac{(n-2)^2\,(\|\beta\|^2)^2\,[3+(n-1)\,\|\beta\|^2](1-\|\beta\|^2)}{[1+(n-1)\,\|\beta\|^2]^3}\right. \\ &\quad \left. + \frac{12(n-2)^2\,\|\beta\|^2}{[1+(n-1)\,\|\beta\|^2]^4}\,(\beta\Sigma\beta')\right\} + O(n^{-2})\,O(n^{-3/2}); \end{aligned}$$

and this simplifies to

$$\begin{aligned} E&\left\{\frac{(n-2)^2\,(\|\mathbf{X}\|^2)^3}{[1+(n-1)\,\|\mathbf{X}\|^2]^2}\right\} \\ &= \frac{(n-2)^2\,(\|\beta\|^2)^3}{[1+(n-1)\,\|\beta\|^2]^2}\left\{1+\frac{1}{n}\left(\frac{1}{\|\beta\|^2}-1\right)\left(1+\frac{2}{1+(n-1)\,\|\beta\|^2}\right)\right. \\ &\quad \left. + \frac{12}{n}\,\frac{\beta\Sigma\beta'}{(\|\beta\|^2)^2}\,\frac{1}{[1+(n-1)\,\|\beta\|^2]^2}\right\} + O(n^{-7/2}). \qquad \text{(A.3)} \end{aligned}$$

Applying Lemma A.1 again with $g(y) = (n-2)\,y/(1+(n-1)\,y)$ and $h(y) = y$, we have that

$$E\left\{\frac{(n-2)\,\|\mathbf{X}\|^2\,\mathbf{X}\beta'}{1+(n-1)\,\|\mathbf{X}\|^2}\right\}$$

$$= \frac{(n-2)(\|\beta\|^2)^2}{1+(n-1)\,\|\beta\|^2} + \frac{1}{n}\left\{\frac{(n-2)\,\|\beta\|^2}{[1+(n-1)\,\|\beta\|^2]^2}(1-\|\beta\|^2)\right.$$

$$+ \frac{1}{2}\left[\frac{4(n-2)}{[1+(n-1)\,\|\beta\|^2]^2}\right.$$

$$\left.\left. + 4\,\|\beta\|^2\left(\frac{-2(n-2)(n-1)}{[1+(n-1)\,\|\beta\|^2]^3}\right)\right](\beta\Sigma\beta')\right\} + O(n^{-1})\,O(n^{-3/2});$$

$$= \frac{(n-2)(\|\beta\|^2)^2}{1+(n-1)\,\|\beta\|^2}\left\{1+\frac{1}{n}\left(\frac{1}{\|\beta\|^2}-1\right)\frac{1}{[1+(n-1)\,\|\beta\|^2]}\right.$$

$$\left. + \frac{2}{n}\frac{\beta\Sigma\beta'}{(\|\beta\|^2)^2}\frac{[1-(n-1)\,\|\beta\|^2]}{[1+(n-1)\,\|\beta\|^2]^2}\right\} + O(n^{-5/2});$$

which after absorbing terms of order $O(n^{-3})$ into $O(n^{-5/2})$ becomes

$$E\left\{\frac{(n-2)\,\|\mathbf{X}\|^2\,\mathbf{X}\beta'}{1+(n-1)\,\|\mathbf{X}\|^2}\right\}$$

$$= \frac{(n-2)(\|\beta\|^2)^2}{1+(n-1)\,\|\beta\|^2}\left\{1+\frac{1}{n}\left(\frac{1}{\|\beta\|^2}-1\right)\frac{1}{[1+(n-1)\,\|\beta\|^2]}\right.$$

$$\left. - \frac{2}{n}\frac{\beta\Sigma\beta'}{(\|\beta\|^2)^2}\frac{1}{[1+(n-1)\,\|\beta\|^2]}\right\} + O(n^{-5/2}). \tag{A.4}$$

Substituting (A.2), (A.3), and (A.4) in (A.1) and expressing in terms of ρ^2 defined in Lemma 2.1 and γ^2 defined in Theorem 3.1, we then obtain

$$R(\delta_3^*, \theta) = \frac{n^2\rho^2}{(n-1)(n-2)}\left\{\frac{(n-2)^2\,\rho^{-6}}{[1+(n-1)\rho^{-2}]^2}\right.$$

$$\times\left[1+\frac{1}{n}(\rho^2-1)\left(1+\frac{2}{1+(n-1)\rho^{-2}}\right)\right.$$

$$\left.\left. + \frac{12\gamma^2}{n[1+(n-1)\rho^{-2}]^2}\right] + O(n^{-7/2})\right\} - \frac{2n\rho^2}{(n-1)}\left\{\left[\frac{(n-2)\rho^{-4}}{1+(n-1)\rho^{-2}}\right]\right.$$

$$\times\left[1+\frac{1}{n}(\rho^2-1)\frac{1}{[1+(n-1)\rho^{-2}]} - \frac{2\gamma^2}{n[1+(n-1)\rho^{-2}]}\right]$$

$$\left. + O(n^{-5/2})\right\} + 1,$$

and this simplifies into

$$R(\delta_3^*, \theta) = \frac{n^2(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)^2} \left\{ 1 + \frac{1}{n}(\rho^2-1) + \frac{2\rho^2(\rho^2-1)}{n(n+\rho^2-1)} + \frac{12\gamma^2\rho^4}{n(n+\rho^2-1)^2} \right\} - \frac{2n(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)} \times \left\{ 1 + \frac{\rho^2(\rho^2-1)}{n(n+\rho^2-1)} - \frac{2\gamma^2\rho^2}{n(n+\rho^2-1)} \right\} + 1 + O(n^{-5/2}).$$

Absorbing $O(n^{-3})$ terms into the $O(n^{-5/2})$ term and regrouping yields

$$R(\delta_3^*, \theta) = \left\{ \frac{n^2(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)^2} - \frac{2n(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)} + 1 \right\} + \left\{ \frac{n^2(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)^2} \frac{2}{n} \frac{\rho^2(\rho^2-1)}{(n+\rho^2-1)} - \frac{2n(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)} \frac{\rho^2(\rho^2-1)}{n(n+\rho^2-1)} + \frac{4n(n-2)}{(n-1)} \frac{1}{(n+\rho^2-1)} \frac{\gamma^2\rho^2}{n(n+\rho^2-1)} \right\} + \left\{ \frac{n^2(n-2)}{(n-1)} \frac{(\rho^2-1)}{n(n+\rho^2-1)^2} \right\} + O(n^{-5/2}).$$

Now, the first term in braces simplifies to $(n+\rho^4)/(n+\rho^2-1)^2 + O(n^{-3})$; the second term in braces simplifies to $4\rho^2\gamma^2/(n+\rho^2-1)^2 + O(n^{-3})$; and the third term in braces is equal to $n(\rho^2-1)/(n+\rho^2-1)^2 - (\rho^2-1)/(n+\rho^2-1)^2 + O(n^{-3})$. Thus,

$$\begin{aligned} R(\delta_3^*, \theta) &= \frac{n\rho^2}{(n+\rho^2-1)^2} + \frac{[\rho^4 + 4\rho^2\gamma^2 - (\rho^2-1)]}{(n+\rho^2-1)^2} + O(n^{-5/2}) \\ &= \frac{\rho^2}{(n+\rho^2-1)} + \frac{[-(\rho^2-1)\rho^2 + \rho^4 + 4\rho^2\gamma^2 - \rho^2 + 1]}{(n+\rho^2-1)^2} + O(n^{-5/2}) \\ &= \frac{\rho^2}{(n+\rho^2-1)} + \frac{(1+4\rho^2\gamma^2)}{(n+\rho^2-1)^2} + O(n^{-5/2}), \end{aligned}$$

and this completes the proof of Theorem 3.1. ∎

Acknowledgments

The author thanks the referee and the editor for their careful reading of the manuscript and for providing helpful suggestions, which improved the presentation, and acknowledges Bowling Green State University for the computing resources.

References

[1] Arnold, B. (1968). Parameter estimation for a multivariate exponential distribution. *J. Amer. Statist. Assoc.* **63** 848–852.

[2] Bemis, B., Bain, L., and Higgins, J. (1972). Estimation and hypothesis testing for the parameters of a bivariate exponential distribution. *J. Amer. Statist. Assoc.* **67** 927–929.

[3] Efron, B. (1967). The two sample problem with censored data. In *Proceedings, Fifth Berkeley Symposium in Math. Stat.*, Vol. IV, pp. 831–853. Prentice-Hall, New York.

[4] Ferguson, T. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York.

[5] Lehmann, E. (1983). *Theory of Point Estimation.* Wiley, New York.

[6] Marshall, A., and Olkin, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62** 30–44.

[7] Peña, E. (1989). *Improved Estimation for an Exponential-Multinomial Distribution with Applications to the Marshall–Olkin MVE Distribution.* Technical Report No. 89-07, Department of Mathematics and Statistics, Bowling Green State University.

[8] Peña, E., and Gupta, A. (1990). Bayes estimation for the Marshall–Olkin exponential distribution. *J. Roy. Statist. Soc. B* **52** 379–389.

[9] Proschan, F., and Sullo, P. (1974). Estimating the parameters of a bivariate exponential distribution in several sampling situations. In *Reliability and Biometry* (Proschan and Serfling, Eds.), pp. 423–440. SIAM, Philadelphia.

[10] Proschan, F., and Sullo, P. (1976). Estimating the parameters of a multivariate exponential distribution. *J. Amer. Statist. Assoc.* **71** 465–472.

[11] Shamseldin, A., and Press, S. (1984). Bayesian parameter and reliability estimation for a bivariate exponential distribution. *J. Econom.* **24** 363–378.